Principal Component Analysis for Raindrops and its Application to the Remote Sensing of Rain

By Jonathan P. Meagher and Ziad S. Haddad*

Jet Propulsion Laboratory, California Institute of Technology, Pasadena CA, USA

(Received November 18, 1997)

SUMMARY

In the problem of inverting remote sensing measurements of rain, current representations of the raindrop size distribution (DSD) suffer crucially from the expedient but unjustified and empirically ill-fitting assumption that the distribution has a known closed-form shape, whether log-normal or Γ-distributed. This paper proposes an approach to avoid such unfounded a priori assumptions entirely. The resulting representation of the rain is then used to derive "forward" formulas for rain remote-sensing algorithms.

KEYWORDS: rain estimation radar drop size distribution

1. INTRODUCTION

The approach currently widely used to establish radar-reflectivity ↔ rain-rate relations from experimental data, and subsequently estimate rainfall from radar measurements, is based on the physical relation between the rain rate R and drop size density function and the radar reflectivity coefficient Z (see e.g. Marshall and Palmer, 1948, and Ryde, 1946). Originally, a simple power law $Z = aR^b$ was assumed, and regression analyses of measured data consisting of simultaneous observations of rain intensities and radar reflectivities were performed, resulting in a plethora of power-law Z-R relations with large variations in the value of the coefficient a and the exponent b (see e.g. Battan, 1973). Other relations were calculated from disdrometer-measured drop size histograms: an analytic form for the drop size distribution was postulated (log-normal or Γ), then the parameters of the distribution were calculated from the data, typically using notoriously biased sample moments. The values of the resulting coefficients a and b still ranged over wide intervals. More serious is the problem that the approach does not guarantee that the parameters are mutually independent, or indeed that they are mutually independent with R (in fact, quite the opposite is true), leading to very serious inconsistencies in the algorithms that use such relations to retrieve rain. One "quick fix" solution would be to eschew DSD-based relations altogether and use only regression-based power laws. However, the problems with the regression relations are much more serious than those with the current DSD approaches (see e.g. Haddad and Rosenfeld, 1997): the integration time required to obtain a sufficiently large set of *simultaneous* samples almost guarantees that the sampled population will not be homogeneous, and the scatter about the mean regression relation produces large uncertainties in the rain retrievals.

A second, equally serious, problem with DSD-based Z-R relations is that statistical tests for goodness of fit have repeatedly failed to support the assumption that the sampled drops are consistent with a Γ- or lognormal distribution. Never close enough to the data when judged by the residual noise, the Γ and lognormal fits are especially bad when large drops occur, i.e. during convective events, and in cases with peaks, convective or stratiform. Recent experiments have shown evidence of drop breakup with peaks near 0.7 mm and 2.5 mm (Keenan, 1997). This basic mismatch between the assumption about the DSD shape and the actual data will produce radar-rain relations that are ill-suited to the type of rain event under study, and that will unavoidably bias the rain estimates.

* Corresponding author: 300-227, JPL, 4800 Oak Grove Drive, Pasadena CA 91109-8099, USA

2. Surface Data

The approach we adopted to solve this problem is based on Karhunen's theorem for stochastic processes. Indeed, the main advantage of parametrising the DSD is to reduce the entire description of the rain (at least of that aspect of the rain that affects the radiometric measurements) to the knowledge of two or (more typically) three values, for example in the case of the Γ-distribution representation (Ulbrich, 1983) two shape parameters μ and Λ representing the normalised variance of the drop diameter (actually equal to $1/(\mu+1)$) and the mean drop size divided by the normalised variance (equal to $1/\Lambda$), and a quantity parameter (proportional to the total liquid mass). However, while the parameters that one finds may give the best Γ- or lognormal-distribution fit, the discrepancy between this fit and the data is in almost every case still too large given the number of sampled drops. An intuitively more direct and representative approach would be to use as DSD variables the numbers of drops of all (sufficiently finely discretised) sizes, spanning the entire spectrum of precipitating liquid drops. For example, one could use as parameters $N_1, \cdots, N_{20}$ representing the "recalibrated" Joss-Waldvogel discretisation, namely N_i = the number of drops (per cubic meter of air) with diameter between D_{i-1} and D_i, and $D_0 = 0$, $D_1 = 0.48$ mm, $D_2 = 0.6$ mm, $D_3 = 0.72$ mm, $D_4 = 0.84$ mm, $D_5 = 0.96$ mm, $D_6 = 1.2$ mm, $D_7 = 1.44$ mm, $D_8 = 1.68$ mm, $D_9 = 1.92$ mm, $D_{10} = 2.16$ mm, $D_{11} = 2.52$ mm, $D_{12} = 2.88$ mm, $D_{13} = 3.24$ mm, $D_{14} = 3.6$ mm, $D_{15} = 3.96$ mm, $D_{16} = 4.44$ mm, $D_{17} = 4.92$ mm, $D_{18} = 5.4$ mm, $D_{19} = 6.0$ mm, $D_{20} = \infty$. The problem would then be that one would end up with 20 parameters N_i to be determined when performing a retrieval, definitely too many variables. Grouping adjacent size bins into single variables (each representing a correspondingly wider range of drop sizes) would be counter-productive since it would drastically increase the error in the resulting radar-rain relations (after all, the reflectivity depends on the 6^{th} power of the drop diameter, so if the error in the latter is tripled, say, by combining three adjacent size bins, the error in the reflectivity gets multiplied at least 18-fold!). To avoid this problem, one would need a more careful method to reduce the information in the size bins into two or three variables. This can indeed be accomplished using the Karhunen-Loève approach. In the case at hand, one needs to calculate the covariances of the variables representing the equivalent mass-per-volume-of-air of the drops in each "high-resolution" drop-size bin, then diagonalise the covariance matrix: the eigenvectors corresponding to the three or four largest eigenvalues would be the (three or four) linear combinations of the bin counts which embody (most of) the description of the given DSD, since their eigenvalues are the largest (recall that the eigenvalues are the covariances themselves, so the eigenvectors corresponding to the largest eigenvalues are the variables that vary most, while the ones corresponding to the smaller eigenvalues are the ones that remain relatively constant). An important additional advantage is that the eigenvectors are automatically uncorrelated, thus allowing one to assume correctly that they vary independently (to first order).

We have performed the Karhunen-Loève principal-component analysis on the Joss-Waldvogel distrometer data collected at sea level near Darwin during two rainy seasons, from November 1988 to March 1990. In order not to give equal weight to all drops, it was necessary to choose a more physical weighting which converts counts into mass and thus gives as much importance to a drop-size bin as the *mass* of liquid it represents indicates. We also needed to make sure that the variables could not ever be negative. These concerns naturally lead to the definition of new variables $L_j = \sqrt{(4/3)\pi(D_j/2)^3 N_j}$, which we shall use instead of the N_j's. Note that the norm-squared $\sum_j L_j^2$ gives the total liquid water content. The sample covariance matrix of the L_j's, computed using the Darwin data, can

then be diagonalised. The new (uncorrelated) variables with the largest variances are

$$N'_1 = 0.13L_7 + 0.31L_8 + 0.45L_9 + 0.46L_{10} + 0.5L_{11} + 0.36L_{12} + 0.25L_{13} \quad (1)$$

$$N'_2 = 0.11L_3 + 0.13L_4 + 0.22L_5 + 0.44L_6 + 0.54L_7 + 0.45L_8 + 0.21L_9 - 0.2L_{11} - 0.25L_{12} - 0.25L_{13} - 0.14L_{14} \quad (2)$$

$$N'_3 = 0.1L_3 + 0.13L_4 + 0.22L_5 + 0.38L_6 + 0.23L_7 - 0.34L_9 - 0.26L_{10} + 0.34L_{12} + 0.54L_{13} + 0.33L_{14} + 0.1L_{15} \quad (3)$$

These coefficients are quite remarkable. The first variable does indeed appear to characterise the larger-drop DSD peak, being the weighted sum of the contributions from those drops whose diameter is about $D_{10} = 2.16$ mm, while the second variable is most sensitive to the smaller-diameter drops around $D_6 = 1.2$ mm. Those are remarkably close to the two independently observed DSD peaks (Keenan, 1997). Table 1 lists the variances of all twenty eigenvariables. Note that the variance of N'_4 is already 10 times smaller than that of N'_1, confirming that most of the characterising information about drop quantity and distribution shape is indeed contained in the first three eigenvariables.

One may thus simplify the description of a particular DSD sample by retaining only the corresponding values of (N'_1, N'_2, N'_3, N'_4) and considering that the values of the higher-order N'_j's are their respective means. This procedure is justified by the fact that the variance of N'_4 (and therefore of N'_j, $j \geq 5$) is quite small. Figure 1 shows an example of an original sample, along with its reconstruction using mean values for the higher-order N'_j's. The truncation error is manifestly quite small. More generally, the effect of the truncation error can be quantified using a χ^2 test, calculating for each sample the statistic

$$\sum_{j=1}^{20} \frac{(L_j - L'_j)^2}{L_j} \quad (4)$$

where $\{L_j\}$ are the observed contributions to the liquid mass in each size bin, and $\{L'_j\}$ are the contributions calculated from (N'_1, N'_2, N'_3, N'_4) and the means of the remaining eigenvariables $N'_5, \cdots, N'_{20}$. Of the 6905 samples from Darwin, a quite respectable 3522 fall within the 95^{th} percentile of the distribution of (4). This contrasts quite favorably with the results obtained when Γ or lognormal fits are made, in which case typically *not one* sample passes the classic goodness-of-fit test.

3. Airborne Measurements

Unfortunately, the Darwin data exhibited flagrantly anomalous behavior for larger rain rates R, namely a sudden jump in the correlation between the width of the DSD and R, when R exceeded 12 mm/hr. Since those data were collected using an instrument which has been shown to be non-stationary, especially when exposed to higher rain rates (Sheppard and Joe, 1994), we restricted the samples used in the principal component analysis above to those producing rain rates below 12 mm/hr. To confirm that the results are essentially still valid at higher rain rates, it was necessary to analyse DSD measurements from other tropical locations using different instruments. We chose to analyse measurements made during the Intense Observation Period of the TOGA Coupled Ocean-Atmosphere Response Experiment (TOGA/COARE). These data were collected in the warm pool of the western equatorial Pacific between November, 1992, and February, 1993 (Lukas et al, 1995), using NCAR's 2-D PMS spectrometer probes

mounted on the NCAR Electra aircraft (Yuter et al, 1995). The data were reduced using a method essentially similar to the one described in Black and Hallett, 1986, then re-sampled according to the Darwin distrometer size bins. To "verify" the decomposition obtained from the Darwin data, we first analysed only those TOGA/COARE samples producing rain rates below 10 mm/hr. The first three eigenvariables were

$$\begin{aligned} N_1'' &= 0.13L_4 + 0.19L_5 + 0.43L_6 \\ &\quad + 0.52L_7 + 0.46L_8 + 0.33L_9 + 0.29L_{10} + 0.23L_{11} + 0.15L_{12} \end{aligned} \tag{5}$$

$$\begin{aligned} N_2'' &= 0.37L_1 + 0.34L_2 + 0.4L_3 + 0.35L_4 + 0.35L_5 + 0.33L_6 \\ &\quad - 0.13L_8 - 0.24L_9 - 0.25L_{10} - 0.25L_{11} - 0.14L_{12} \end{aligned} \tag{6}$$

$$\begin{aligned} N_3'' &= 0.5L_1 + 0.32L_2 + 0.28L_3 - 0.33L_6 - 0.27L_7 \\ &\quad + 0.2L_9 + 0.33L_{10} + 0.38L_{11} + 0.28L_{12} + 0.12L_{13} \end{aligned} \tag{7}$$

These coefficients are remarkably similar to those obtained for the Darwin data. The first two variables are concentrated near the same two drop-diameter peaks, and the third is a three-humped window of roughly the same shape as in the Darwin case.

Encouraged by this comparison, we decided to use the entire TOGA/COARE data set to derive expressions for N_j' which would be valid at all rain rates, for all types of rain. Table 2 shows all the entries of the change-of-basis matrix expressing $\{N_j'\}$ in terms of $\{L_j\}$. The means and variances (eigenvalues) of the new (eigen)variables are given in table 3. They are quite similar to those obtained with the Darwin data: the first variable is again concentrated around the larger drop diameters, the second represents a difference between smaller and larger drop contributions, and the third is a three-humped window of the same shape as before. Finally, note that, because the matrix of change of basis is orthogonal, we still have

$$\sum_j N_j'^2 = \text{the total liquid water content} \tag{8}$$

In particular, the variances in table 3 (and 1) are in units of grams per cubic meter. Evidently, the first three eigenvariables embody most of the quantitative and qualitative information about a DSD sample.

4. Horizontal and Vertical Variability

Because the PMS spectrometers were mounted on a platform flying at an approximately constant speed, one can readily use the COARE data to estimate the spatial variability of the DSD eigenvariables. One measure of variability is particularly useful in our case: the absolute m.s. variation v_δ defined for a stationary random process $N'(t)$ simply as

$$v_\delta = \mathcal{E}\{(N'(t) - N'(t+\delta))^2\} \tag{9}$$

Table 4 shows the values for v_δ obtained from the COARE data for the eigenvariables N_1', N_2', N_3' and N_4' (the spatial auto-correlation of the higher-order N''s are not shown because their standard deviation over the *entire* data is already negligibly small, as is evident in table 3). These variations can be directly compared to the r.m.s. variation of the total liquid water $\sum L_j^2 = \sum N_j'^2$ included in table 4. The sample size was insufficient to calculate correlations beyond $\delta = 8$ km with much confidence. The results confirm that the spatial variation of N_j' for $j \geq 2$ remains quite small indeed.

Finally, table 5 shows the covariances between the values of the N'_j's for the Darwin samples (again with $R \leq 12$ mm/hr). The fact that the off-diagonal entries are quite small relative to the diagonal variances implies that our principal component representation, derived form the COARE data sampled aloft, retains its first-order independence for ground DSD samples.

5. Radiometric Relations

Weather radars can measure the effective reflectivity Z_e of rainfall quite accurately (see, e.g., Battan, 1973). At the higher frequencies typical of planned spaceborne designs, the measured reflectivity is lower than the true Z_e because of the attenuation $\int_\gamma k$ accumulated along the propagation path γ, where k is the attenuation coefficient. The problem of estimating the rain rate R given attenuated reflectivity measurements can be expressed using Z-R and k-R relations. More recently, the specific polarisation propagation differential phase shift Φ_{DP} and the differential reflectivity Z_{DR} have also been suggested for their correlation with R and their relatively weak dependence on drop size. Naturally, there are numerous Z-R, k-R, Z_{DR}-R and Φ_{DP}-R relations for any given frequency (see, e.g., Olsen, 1978), ultimately depending on the shape of the drop size distribution (or at least, in the case of Φ_{DP}, on the mean drop diameter), and on other environmental factors. Since an inappropriate relation could lead to serious errors in the retrieved rainfall (Haddad, 1995), it is particularly useful to have relations that are explicitly parametrized by the DSD: one would then try to determine the appropriate parameters either from one's data or from ancillary observations. A parametrization which uses the principal component analysis above would be particularly useful (and unique) because it would make no a priori assumption about the form of the DSD, and it would allow one to assume constant however many DSD variables one must without committing any correlation-induced inconsistencies and while quantifying the r.m.s uncertainty which the constancy assumption will have introduced.

To obtain DSD-based relations between Z_e and R, we assigned to the vector (N'_2, N'_3) regularly-spaced discrete values within two standard deviations of the means of each of the variables: in each case, we then used a Mie-scattering model to compute Z_e exactly as N'_1 (hence R) varied in the range $[0.47 - 3 \times 0.14, 0.47 + 3 \times 0.14]$ (i.e. within three standard deviations of the mean, see table 3), assuming that the temperature varied between 275 K and 290 K, and letting $(N'_4, \cdots, N'_{20})$ vary within two standard deviations of their respective means. The power law minimising the sum of the mean-squared distances from the Mie-calculated reflectivities was then calculated for each pair (N'_2, N'_3). The resulting Z_e-R power-law relations

$$Z_e = a(N'_2, N'_3) R(N'_1; N'_2, N'_3)^{b(N'_2, N'_3)} \tag{10}$$

for the Tropical Rainfall Measuring Mission's 13.8 Ghz frequency are given in table 6. To illustrate the validity of (10), figure 2 shows the Mie and approximate Z_e R curves, when $N'_2 = 4.61$ and $N'_3 = 0.78$. Finally, table 7 gives the Rayleigh relations, which apply for typical ground-based radars. Note that the coefficients do vary quite significantly with N'_2. On the other hand, given a distrometer measurement, one can calculate the value of N'_2 that should be used in retrieving R.

Relations for the microwave attenuation, differential reflectivity, and differential phase can be obtained in the same manner. We intend to use these relations (means and variances) to derive a stochastic filter to estimate the rain means and r.m.s. uncertainties from radar measurements.

6. CONCLUSIONS

The first conclusion one must draw from this application of the principal component analysis to binned raindrop sizes is that the resulting description of drop size distributions is quite robust, producing variables which are essentially mutually uncorrelated even when the correlations are calculated from DSD populations sampled at different times, in different locations and with different instruments.

Equally important, since the results obtained using a ground distrometer do not differ significantly from those obtained using airborne probes, in spite of the vastly different measurement uncertainties, one must conclude that the *joint* statistics of drop sizes do not differ significantly in altitude and at the surface. This is particularly useful in the application of remote sensing to estimate precipitation, since it implies that precipitation can be modeled using the same set of variables at all altitudes.

Moreover, most of these descriptor variables can be assumed constant spatially, since a) the variances of all but the first four variables are indeed negligible, and b) the horizontal autocorrelation estimated from the airborne measurements shows that the all but the first couple of variables vary little spatially. This allows one to reduce the number of unknowns in one's model, without committing the classic inconsistency of assuming one variable constant and another spatially-varying when the two are significantly correlated.

We are currently applying these results to various rain retrieval procedures, using ground, airborne, and, soon, spaceborne radar measurements of rain.

ACKNOWLEDGEMENT

We wish to thank Robert A. Black for graciously sharing the PMS probe data. This work was performed at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration.

REFERENCES

Atlas, D., Ulbrich, C.W., Marks, F.D., Willis, P.T., Samsury, C.E., and Black, R.A. 1996 Tropical rain: microphysics and radar properties, part II – altitude dependence. *submitted to J. Appl. Meteor.*

Battan L.J. 1973 *Radar Observations of the Atmosphere*, The University of Chicago Press

Black, R.A., and Hallett, J. 1986 Observations of the distribution of ice in hurricanes. *J. Atmos. Sci.* **43**, 802-822

Haddad, Z.S., Jameson, A.R., Durden, S.L., and Im, E. 1995 Coupled Z–R and k–R relations and the resulting ambiguities in the determination of the vertical distribution of rain from the radar backscatter and the integrated attenuation. *J. Appl. Meteor.* **34**, 2680–2688

Haddad, Z.S., and Rosenfeld, D. 1997 Optimality of Empirical Z–R relations. *Quart. J. Roy. Met. Soc.* **123**, 1283-1293

Hitschfeld, W., and Bordan, J. 1954 Errors inherent in the radar measurements of rainfall at attenuating wavelengths. *J. Meteor.* **11**, 58–67

Keenan, T.D. 1997 *Personal communication*

Lukas, R., Webster, P.J., Ji, M., and Leetmaa, A. 1995 The large-scale context for the TOGA coupled ocean-atmosphere response experiment. *Meteor. Atmos. Phys.* **56**, 3–16

Marshall, J.S., and Palmer, W.M.K. 1948 The distribution of rain drops with size. *J. Meteor.* **5**, 165-166

Olsen, R., Rogers, D., and Hodge, D. 1978 The aR^b relation in the calculation of rain attenuation. *IEEE Trans. Ant. Prop.* **26**, 318-329

Ryde, J.W. 1946 Attenuation of centimeter radio waves and the echo intensities resulting from atmospheric phenomena. *I.E.E.E. J. pt. 3A* **93**, 101-103

Sheppard, B.E., and Joe, P.I. 1994 Comparison of raindrop size distribution measurements by a Joss-Waldvogel distrometer, a PMS 2DG spectrometer, and a POSS Doppler radar. *J. Atmos. Ocean. Technol.* **11**, 874-887

Tokay, A., and Short, D.A. 1996 Evidence from tropical raindrop spectra of the origin of rain from stratiform and convective clouds. *J. Appl. Meteor.* **35**, 355-371

Ulbrich, C.W. 1983 Natural variations in the analytical form of the raindrop size distribution, on computer simulations of dual-measurement radar methods. *J. Climate Appl. Meteor.* **22**, 1764–1775

Yuter, S., Houze, R.A., Smull, B.F., Marks, F.D., Daugherty, J.R., and Brodzik, S.R. 1995 TOGA COARE aircraft mission summary images - An electronic atlas. *Bull. Amer. Meteor. Soc.* **76**, 319-328

Figure captions

Figure 1: Sample "before/after" histogram illustrating the goodness-of-fit of the "truncated" DSD representation (this sample is from the Darwin data, and the truncation assumed $N'_5, \cdots, N'_{19}$ constant, equal to their observed means)

Figure 2: Sample radar-rain relation at 13.8 GHz, when $(N'_2, N'_3) = (4.61, 0.78)$

N'_1	N'_2	N'_3	N'_4	N'_5	N'_6	N'_7	N'_8	N'_9	N'_{10}	N'_{11}	N'_{12}	N'_{13}	N'_{14}	N'_{15}	N'_{16}	N'_{17}	N'_{18}	N'_{19}
21.3	18.7	6.0	2.2	2.0	1.8	1.5	1.1	0.9	0.7	0.4	0.34	0.28	0.25	0.16	0.15	0.09	0.08	0.06

TABLE 1. Variances ($\times 10^3$) of the Darwin eigenvariables for rain rates under 12 mm/hr

	L_1	L_2	L_3	L_4	L_5	L_6	L_7	L_8	L_9	L_{10}	L_{11}	L_{12}	L_{13}	L_{14}	L_{15}	L_{16}	L_{17}	L_{18}	L_{19}	L_{20}
N'_1	0.20	0.19	0.23	0.27	0.31	0.50	0.49	0.37	0.21	0.14	0.09	0.04	0.02	0.02	0.01	0.01	0.01	0.01	0.00	0.00
N'_2	0.35	0.29	0.32	0.24	0.22	0.10	-0.17	-0.35	-0.37	-0.37	-0.33	-0.19	-0.08	-0.04	-0.02	-0.01	-0.01	-0.01	-0.01	0.00
N'_3	0.48	0.30	0.26	0.04	-0.05	-0.33	-0.32	-0.06	0.18	0.32	0.39	0.29	0.14	0.07	0.04	0.02	0.01	0.01	0.01	0.02
N'_4	-0.32	-0.12	-0.02	0.19	0.22	0.30	-0.08	-0.36	-0.29	-0.01	0.25	0.52	0.28	0.20	0.13	0.09	0.05	0.03	0.03	0.03
N'_5	-0.26	-0.05	0.07	0.22	0.21	0.13	-0.24	-0.24	0.07	0.33	0.37	-0.33	-0.35	-0.31	-0.21	-0.17	-0.12	-0.08	-0.10	-0.15
N'_6	0.06	0.02	0.00	-0.06	-0.05	-0.01	0.08	0.05	-0.04	-0.09	-0.16	0.69	-0.44	-0.35	-0.23	-0.21	-0.14	-0.09	-0.08	-0.17
N'_7	-0.29	-0.01	0.15	0.27	0.19	-0.09	-0.31	0.01	0.49	0.19	-0.61	0.14	0.04	0.11	0.08	0.05	0.03	0.00	0.00	-0.02
N'_8	0.00	-0.01	-0.02	0.02	0.02	0.01	-0.02	-0.02	-0.07	0.14	0.00	0.06	-0.52	-0.06	0.17	0.24	0.21	0.15	0.27	0.68
N'_9	-0.13	0.01	0.09	0.11	0.06	-0.09	-0.11	0.08	0.43	-0.73	0.37	0.02	-0.20	0.04	0.15	0.12	0.04	-0.01	-0.01	-0.04
N'_{10}	-0.05	-0.01	0.00	0.05	0.04	0.00	-0.05	0.00	0.16	-0.19	0.02	0.02	0.42	-0.26	-0.44	-0.35	-0.15	0.01	0.12	0.58
N'_{11}	0.36	-0.07	-0.32	-0.26	-0.07	0.42	0.01	-0.54	0.47	0.00	-0.07	-0.01	-0.04	0.03	0.06	-0.02	-0.03	-0.01	0.03	-0.01
N'_{12}	-0.01	0.00	0.00	0.01	0.00	-0.02	0.01	0.03	-0.04	0.01	0.02	-0.02	-0.26	0.63	0.14	-0.52	-0.45	-0.02	-0.10	0.19
N'_{13}	-0.01	0.00	0.00	0.02	0.01	-0.01	-0.03	0.02	0.01	-0.02	0.01	-0.01	-0.04	0.05	-0.09	-0.19	-0.04	0.55	0.73	-0.32
N'_{14}	-0.01	0.00	0.01	0.01	0.00	-0.01	-0.01	0.03	-0.02	0.02	-0.02	-0.01	0.15	-0.47	0.65	0.03	-0.57	0.04	0.08	0.03
N'_{15}	-0.23	0.22	0.44	-0.53	-0.23	0.46	-0.35	0.18	-0.02	-0.01	0.00	0.00	0.00	0.01	-0.01	0.00	0.00	0.01	0.00	0.02
N'_{16}	-0.24	0.15	0.35	-0.08	-0.16	-0.20	0.45	-0.36	0.09	0.04	0.00	-0.02	0.00	0.02	0.06	-0.06	0.04	-0.48	0.38	-0.02
N'_{17}	0.17	-0.10	-0.27	0.06	0.12	0.15	-0.35	0.30	-0.10	-0.02	0.00	0.00	0.00	0.05	0.01	0.01	-0.01	-0.65	0.44	-0.05
N'_{18}	-0.02	0.01	0.03	-0.01	-0.01	-0.02	0.05	-0.04	0.02	0.01	0.00	0.00	-0.05	0.18	-0.42	0.64	-0.61	0.01	0.08	0.02
N'_{19}	-0.25	0.82	-0.50	0.09	-0.07	0.02	-0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.01	0.00
N'_{20}	-0.03	0.09	-0.05	-0.57	0.78	-0.23	0.05	-0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

TABLE 2. Change-of-basis matrix (from the COARE data)

	N'_1	N'_2	N'_3	N'_4	N'_5	N'_6	N'_7	N'_8	N'_9	N'_{10}	N'_{11}	N'_{12}	N'_{13}	N'_{14}	N'_{15}	N'_{16}	N'_{17}	N'_{18}	N'_{19}	N'_{20}
mean $\times 10$	4.7	0.26	0.74	0.13	0.16	0.01	0.03	0.04	0.01	0.02	0.01	0.008	0.004	0.001	0.04	0.01	0.06	0.003	0.006	0.009
variance $\times 10^3$	20.3	15.5	5.2	2.1	1.1	0.8	0.7	0.6	0.5	0.5	0.37	0.34	0.28	0.25	0.09	0.15	0.17	0.18	0.002	0.003

TABLE 3. Means and Variances of the COARE eigenvariables

	1 km	2 km	3 km	4 km	5 km	6 km	7 km	8 km
$\sum N'^2_j$	0.031	0.046	0.047	0.038	0.046	0.035	0.046	0.05
N'_1	0.018	0.027	0.032	0.036	0.038	0.04	0.042	0.044
N'_2	0.01	0.015	0.017	0.019	0.021	0.022	0.023	0.024
N'_3	0.005	0.008	0.009	0.01	0.01	0.01	0.01	0.011
N'_4	0.004	0.004	0.004	0.004	0.004	0.004	0.005	0.005

TABLE 4. r.m.s. variation of the total liquid mass, and m.s. variation of the first four eigenvariables, in g/m^3 - the mean total liquid $\sum N'^2_j$ is 0.16 g/m^3

	N'_1	N'_2	N'_3	N'_4	N'_5	N'_6	N'_7	N'_8	N'_9	N'_{10}	N'_{11}	N'_{12}	N'_{13}	N'_{14}	N'_{15}	N'_{16}	N'_{17}	N'_{18}	N'_{19}
N'_1	17.46	5.62	-2.90	-2.88	-0.45	-0.58	-0.73	0.20	-0.66	0.21	-0.19	-0.07	0.03	0.06	-0.38	-0.14	0.16	0.09	-0.07
N'_2	5.62	14.25	8.54	0.92	0.13	0.48	-0.42	-1.15	-0.35	0.66	0.56	0.18	0.13	-0.04	0.31	0.24	-0.20	0.01	-0.08
N'_3	-2.90	8.54	11.65	4.76	-0.01	0.91	-0.01	-1.46	0.17	0.67	0.53	0.27	0.14	-0.09	0.38	0.20	-0.21	-0.03	0.05
N'_4	-2.88	0.92	4.76	7.02	1.66	0.64	1.01	-1.54	0.55	0.46	0.39	0.39	0.21	-0.15	0.07	0.04	-0.15	-0.05	-0.04
N'_5	-0.45	0.13	-0.01	1.66	3.69	0.39	-0.27	-1.33	0.03	0.28	-0.39	0.37	0.11	-0.10	0.00	0.02	-0.15	-0.09	0.03
N'_6	-0.58	0.48	0.91	0.64	0.39	2.53	-0.05	-0.88	0.18	0.03	0.11	0.31	0.04	-0.05	0.03	-0.07	-0.07	-0.09	-0.01
N'_7	-0.73	-0.42	-0.01	1.01	-0.27	-0.05	2.20	-0.04	-0.12	0.05	-0.29	-0.07	0.01	-0.02	0.06	-0.09	0.06	0.05	0.02
N'_8	0.20	-1.15	-1.46	-1.54	-1.33	-0.88	-0.04	1.40	-0.12	-0.88	0.03	0.02	-0.10	-0.02	-0.03	-0.06	0.06	-0.03	0.00
N'_9	-0.66	-0.35	0.17	0.55	0.03	0.18	-0.12	-0.12	1.73	-0.06	-0.01	-0.12	0.06	0.10	0.02	-0.07	0.05	0.02	0.01
N'_{10}	0.21	0.66	0.67	0.46	0.28	0.03	0.05	-0.88	-0.06	1.08	-0.01	-0.52	0.01	0.09	0.01	0.10	-0.01	0.04	-0.01
N'_{11}	-0.19	0.56	0.53	0.39	-0.39	0.11	-0.29	0.03	-0.01	-0.01	1.05	0.01	0.02	-0.04	-0.02	-0.10	0.10	0.00	0.00
N'_{12}	-0.07	0.18	0.27	0.39	0.37	0.31	-0.07	0.02	-0.12	-0.52	0.01	1.02	-0.03	-0.39	0.00	-0.02	-0.08	0.13	0.01
N'_{13}	0.03	0.13	0.14	0.21	0.11	0.04	0.01	-0.10	0.06	0.01	0.02	-0.03	0.10	0.10	0.01	-0.01	0.02	0.01	0.00
N'_{14}	0.06	-0.04	-0.09	-0.15	-0.10	-0.05	-0.02	-0.02	0.10	0.09	-0.04	-0.39	0.10	0.83	0.01	0.01	0.04	-0.25	0.00
N'_{15}	-0.38	0.31	0.38	0.07	0.00	0.03	0.06	-0.03	0.02	0.01	-0.02	0.00	0.01	0.01	0.34	-0.10	0.08	-0.01	-0.02
N'_{16}	-0.14	0.24	0.20	0.04	0.02	-0.07	-0.09	-0.06	-0.07	0.10	-0.10	-0.02	-0.01	0.01	-0.10	0.42	-0.25	0.02	-0.01
N'_{17}	0.16	-0.20	-0.21	-0.15	-0.15	-0.07	0.06	0.06	0.05	-0.01	0.10	-0.08	0.02	0.04	0.08	-0.25	0.30	0.00	0.01
N'_{18}	0.09	0.01	-0.03	-0.05	-0.09	-0.09	0.05	-0.03	0.02	0.04	0.00	0.13	0.01	-0.25	-0.01	0.02	0.00	0.49	0.01
N'_{19}	-0.07	-0.08	0.05	-0.04	0.03	-0.01	0.02	0.00	0.01	-0.01	0.00	0.01	0.00	0.00	-0.02	-0.01	0.01	0.01	0.13

TABLE 5. Covariance matrix for the Darwin data in the new (COARE-derived) representation

N'_2	$N'_3 = -8.7$	$N'_3 = -5.5$	$N'_3 = -2.4$	$N'_3 = 0.78$	$N'_3 = 3.9$
-11.72	(2051.59, 0.746)	(1733.93, 0.780)	(1333.23, 0.843)	(927.96, 0.936)	(602.07, 1.049)
-6.28	(1696.19, 0.750)	(1527.62, 0.761)	(1244.96, 0.804)	(961.15, 0.863)	(677.43, 0.950)
-0.83	(973.98, 0.862)	(843.09, 0.887)	(684.85, 0.930)	(531.06, 0.988)	(397.17, 1.056)
4.61	(301.49, 1.162)	(305.34, 1.135)	(258.11, 1.166)	(221.48, 1.191)	(204.18, 1.193)
10.05	(135.47, 1.318)	(147.91, 1.268)	(196.78, 1.153)	(185.66, 1.156)	(246.41, 1.050)

TABLE 6. (a, b) pairs for the reflectivity–rain relation $Z = aR^b$ at 13.8 Ghz

N'_2	$N'_3 = -8.7$	$N'_3 = -5.5$	$N'_3 = -2.4$	$N'_3 = 0.78$	$N'_3 = 3.9$
-11.72	(731.11, 0.885)	(659.51, 0.912)	(559.15, 0.959)	(454.11, 1.019)	(365.05, 1.079)
-6.28	(622.25, 0.921)	(588.79, 0.934)	(540.75, 0.957)	(482.39, 0.987)	(414.99, 1.030)
-0.83	(466.94, 0.996)	(466.94, 0.994)	(457.77, 0.997)	(438.74, 1.007)	(414.43, 1.021)
4.61	(300.32, 1.120)	(346.72, 1.073)	(371.67, 1.049)	(392.18, 1.031)	(406.79, 1.018)
10.05	(248.93, 1.157)	(278.62, 1.124)	(327.84, 1.075)	(375.09, 1.034)	(427.78, 0.994)

TABLE 7. (a, b) pairs for the Rayleigh reflectivity–rain relation $Z = aR^b$

figure 1

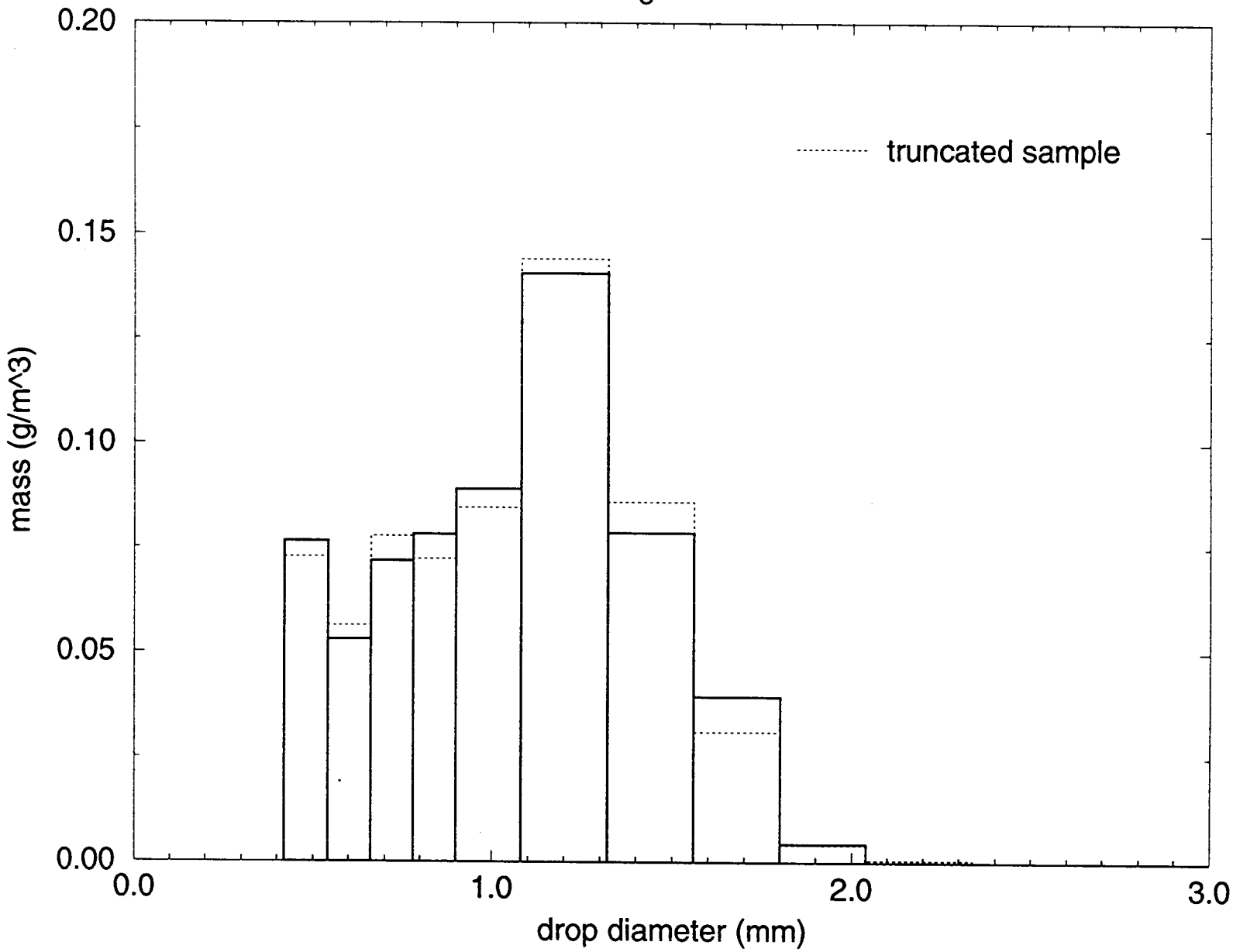
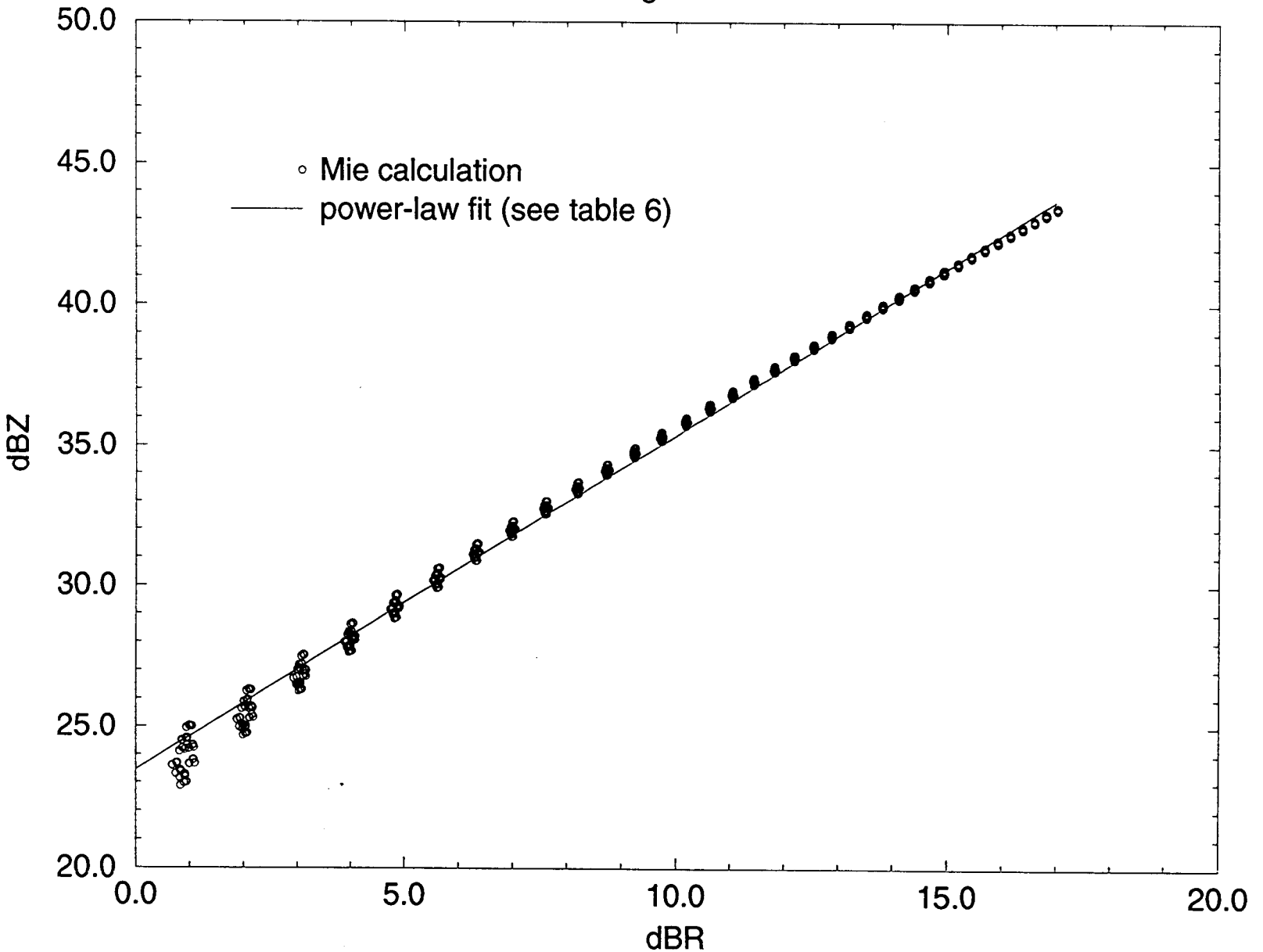

0.20
0.15
0.10
0.05
0.00
mass (g/m^3)
truncated sample
0.0
1.0
2.0
3.0
drop diameter (mm)

figure 2


50.0
45.0
40.0
35.0
30.0
25.0
20.0
dBZ
0.0
5.0
10.0
15.0
20.0
dBR
Mie calculation
power-law fit (see table 6)